Convergence Criteria for Transversal Equalizers

By D. W. LYTLE*

(Manuscript received March 22, 1968)

Two basic problems in the equalization of data channels using pulse-amplitude modulation are considered. The first of these is to determine just what pulses can be equalized and what the general equalization solution is when using a transversal filter. The second problem is to determine if a simple iterative search routine will converge to a solution if its exists.

The unequalized channel impulse response is represented by a polynomial whose coefficients are the sample values of the impulse response. If no roots of this polynomial lie on the unit circle, the channel can be equalized. The transversal filter which equalizes the pulse has tapweight values given by weighted sums of powers of the polynomial roots.

Various necessary and sufficient conditions for iterative convergence are developed. Iterative convergence can be guaranteed if the proper linear weighting of the output sample errors is used in adjusting the tap-weights.

I. INTRODUCTION

This paper is concerned with certain aspects of the automatic equalization of low-noise, linear channels which are to be used for multilevel pulse-amplitude modulated (PAM) signals. The purpose of the equalizer is to compensate for the channel transfer characteristics in such a way that the over-all impulse response of the channel is a Nyquist-I type of pulse,[1, 2, 3] that is, as is illustrated in Fig. 1, a pulse with a central peak and uniformly spaced zeros with period T. If such an impulse response is achieved, a sequence of amplitude modulated impulses with period T can be transmitted and the sequence of amplitudes can be recovered at the receiver by simply sampling in synchronism with period T.

Certain obvious questions such as how to achieve synchronism and

* University of Washington, Seattle.

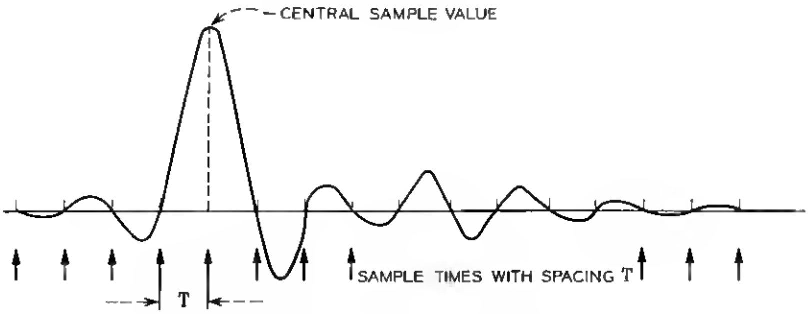

Fig. 1 — A Nyquist-I pulse with central sample and zeros at periodic sample times.

the effects of sampling jitter, nonlinear distortion, and additive noise, although of great importance, are neglected in this presentation in order to concentrate on the methods of adjusting the equalizer. Thus, we assume a perfectly synchronized, noiseless, linear channel with an ideal sampler. Some further constraints which should simplify and clarify the presentation are as follows. We specify that the equalizer is to be a transversal (tapped delay-line) filter with tap weights $\{\alpha_k\}$ which can be adjusted. If the input to this filter, illustrated in Fig. 2, is $\beta(t)$, then the output, $\gamma(t)$, is

$$\gamma(t) = \sum_{j=-n}^{N} \alpha_j \beta(t + jT). \tag{1}$$

Notice the tap-weight numbering convention and the treatment of the delay-line as being composed of negative as well as positive delay. These conventions will simplify the notation in future derivations.

Although the equalizer may be placed at many points within the communication system, for convenience we will consider it to be the final component other than the final sampler. Thus, tbe objective is

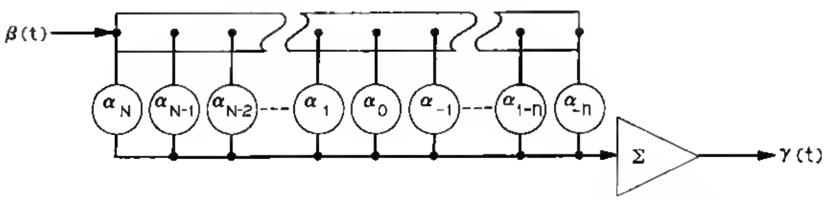

Fig. 2 — A transversal filter with input $\beta(t)$, tap weights α_j, and output $\gamma(t)$.

to adjust the tap weights so that the output, $\gamma(t)$, is a Nyquist-I pulse when the input, $\beta(t)$, is the impulse response of the system before equalization.

As a final simplification we assume that the tap-weight adjustment is to be carried out when the impulse response is available at the input. This "training period" assumption avoids the added complexities of extracting channel characteristics when data are being transmitted.[4] However, the properties to be developed can be readily extended to continuously adaptive equalization.

II. TIMING AND THE ROOTS OF THE IMPULSE RESPONSE

Since the output of the equalizer is to be sampled, we need consider only its effect on the output sample values $\{\gamma_k\}$. Ideally, with an impulse applied to the channel input, one of the output samples, say γ_0, should have unit amplitude while all the others are zero. The tap weights are to be adjusted in order to approach this goal. An additional parameter which will affect the equalization is the timing.

Let the sample values of the impulse response at the equalizer input be the set $\{\beta_j\}$ where

$$\beta_j = \beta(jT) \tag{2}$$

and

$$\beta(t) = h(t + \tau). \tag{3}$$

Equation (3) is to indicate that the sampling times are arbitrary. This is, if $h(t)$ is the channel response to an impulse applied at $t = 0$, then the sample set $\{\beta_j\}$ is a function of the factor τ. Notice that our assumption of perfect synchronization means that the periodicity factor, T, in equation (2) is the proper value. But it does not imply that τ is prescribed. We shall see that the operation of the equalizer depends very strongly upon the value of τ.

This impulse response is to be equalized by the transversal filter with tap weights $\{\alpha_k\}$. The transversal filter output sample set is $\{\gamma_k\}$ where

$$\gamma_k = \sum_j \alpha_j \beta_{k+j} . \tag{4}$$

The sample set $\{\beta_j\}$ will be considered finite in extent, that is,

$$\beta_j = 0 \quad \text{for} \quad j < -m \quad \text{and} \quad j > M. \tag{5}$$

This is a reasonable approximation for any actual channel.

This sample set may be treated as the coefficients of a polynomial $B(z)$.

$$B(z) = \beta_M z^{M+m} + \beta_{M-1} z^{M+m-1} + \cdots + \beta_0 z^m + \cdots + \beta_{-m+1} z + \beta_{-m} . \quad (6)$$

In factored form $B(z)$ may be written as

$$B(z) = \beta_M (z - \theta_1)(z - \theta_2) \cdots (z - \theta_I)(z - \phi_1)(z - \phi_2) \cdots (z - \phi_\Omega) \quad (7)$$

where the roots inside the unit-circle are denoted by the θ values and the roots outside by the ϕ values.

$$| \theta | < 1 < | \phi | \quad (8)$$

$$I + \Omega = M + m. \quad (9)$$

As we have noted, the sample set $\{\beta_j\}$ is a function of the factor τ. Thus, for any particular channel, the roots of equation (7) will wander as τ is varied. Each root will wander on some cyclic path which has a period T. That is

$$\theta_i(\tau + nT) = \theta_i(\tau) \quad (10)$$

This is illustrated in Fig. 3 where a pulse shape is shown and in Fig. 4 where the root loci are shown for variations in τ. Notice that as τ increases from 0 to T, at least one of the roots, regardless of the pulse shape, will cross the unit circle.

Perhaps the periodicity of the roots can be better understood if the sampling is thought of as multiplication by a comb of impulses with spacing T. Each impulse has associated with it a power of z. For example, in equation (6), we see that the impulse yielding the earliest nonzero sample (β_{-m}) is associated with the zero power of z, the next impulse yielding β_{-m+1} is associated with the first power, and so on. As the comb is moved relative to the pulse, $\beta(t)$, the impulses produce different samples and when moved a whole period T, the comb will reproduce the original samples again. However, each sample would be paired with a one-higher or one-lower power of z than previously. Thus, for example, if the comb were shifted by T so that the powers of z were one higher, the factorization of equation (7) would be obtained with the same roots except for an additional root at $z = 0$ since the original polynomial is multiplied by the first power of z. As the comb is shifted along, the additional root, which must eventually go to $z = 0$, comes in from $z = \infty$ and at some particular shift crosses the unit circle.

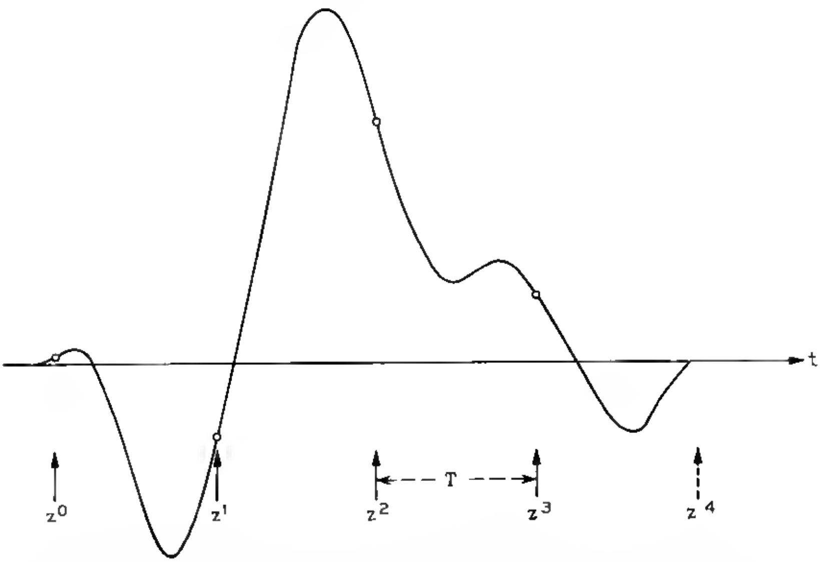

Fig. 3 — A pulse and four nonzero sample positions which yield a polynomial with roots marked 1 in Fig. 4. Nine additional sets of roots are obtained by moving the positions above to the right in increments of $T/10$.

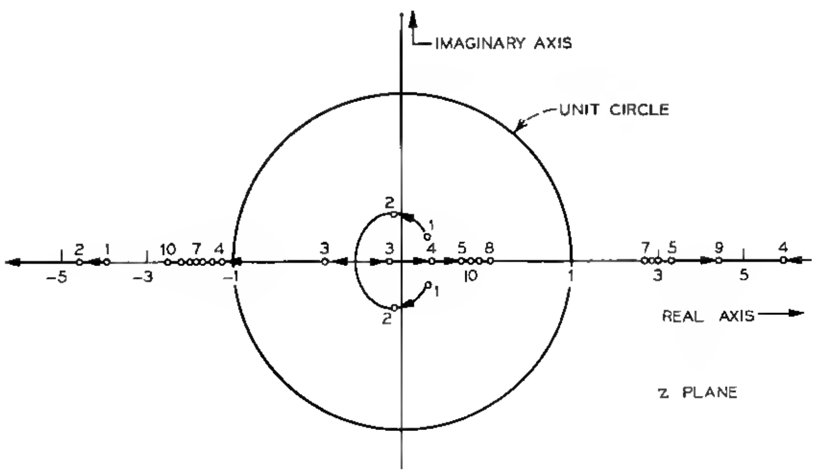

Fig. 4 — Root loci for pulse of Fig. 3. Scale inside circle is magnified four times.

III. IDEAL TAP-WEIGHT ADJUSTMENT

Let us now consider the problem of forcing the γ_k to zero. The desired conditions are

$$\gamma_0 = \sum_j \alpha_j \beta_j = 1 \tag{11}$$

and

$$\gamma_k = \sum_j \alpha_j \beta_{j+k} = 0 \qquad k \neq 0. \tag{12}$$

Equations (11) and (12) in the form of a matrix array is as follows. We shall assume temporarily that the number of taps is infinite.

$$\begin{bmatrix} & & & \vdots & & & \\ \cdots\ 0\ 0\ \beta_M \cdots \cdots \beta_2\ \beta_1\ \beta_0\ \beta_{-1}\ \beta_{-2} \cdots \cdots \beta_{-m}\ 0 & & & & & & \\ \quad 0\ 0\ \beta_M \cdots \beta_3\ \beta_2\ \beta_1\ \beta_0\ \beta_{-1}\ \beta_{-2} \cdots \cdots \beta_{-m}\ 0 \cdots & & & & & & \\ \qquad 0\ \beta_M \cdots \beta_3\ \beta_2\ \beta_1\ \beta_0\ \beta_{-1}\ \beta_{-2} \cdots \cdots \beta_{-m}\ 0 \cdots & & & & & & \\ \qquad\quad 0\ \beta_M \cdots \beta_3\ \beta_2\ \beta_1\ \beta_0\ \beta_{-1}\ \beta_{-2} \cdots \cdots \beta_{-m} \cdots & & & & & & \\ & & & \vdots & & & \end{bmatrix} \cdot \begin{bmatrix} \vdots \\ \alpha_2 \\ \alpha_1 \\ \alpha_0 \\ \alpha_{-1} \\ \alpha_{-2} \\ \alpha_{-3} \\ \vdots \end{bmatrix} = \begin{bmatrix} \vdots \\ 0 \\ 0 \\ 1 \\ 0 \\ 0 \\ 0 \\ \vdots \end{bmatrix}. \tag{13}$$

Now let us consider a typical equation of the form,

$$\sum_{j=-m-k}^{M-k} \alpha_j \beta_{k+j} = 0 \tag{14}$$

where we will take either the case where $k < -m$ or $k > M$. If these conditions on k are satisfied, then equation (14) involves α values with only positive subscripts or only negative subscripts. Equation (14) is a linear homogeneous difference equation in the variable α_j. Such

equations have solutions of the form

$$\alpha_j = z^j. \tag{15}$$

Substitution of equation (15) into equation (14) yields

$$\begin{aligned} &\beta_M z^{M-k} + \beta_{M-1} z^{M-k-1} + \cdots + \beta_1 z^{1-k} \\ &\qquad + \beta_0 z^{-k} + \beta_{-1} z^{-1-k} + \cdots + \beta_{-m} z^{-m-k} \\ &\quad = z^{-m-k}[\beta_M z^{M+m} + \beta_{M-1} z^{M+m-1} + \cdots + \beta_0 z^m + \beta_{-m}] \\ &\quad = z^{-m-k} B(z) = 0. \end{aligned} \tag{16}$$

Equation (16), and hence equation (14), will be satisfied for z equal to one of the roots of the impulse response polynomial $B(z)$. And since we have a linear homogeneous equation, any linear combination of solutions is a solution. Thus, the general solution is

$$\alpha_j = C_1\theta_1^j + C_2\theta_2^j + \cdots + C_I\theta_I^j + D_1\phi_1^j + \cdots + D_\Omega\phi_\Omega^j \tag{17}$$

The constants, $C_1, C_2, \ldots, C_I, D_1, D_2, \ldots, D_\Omega$, are arbitrary and will be adjusted to meet the boundary conditions. One boundary condition which should be imposed is the following. Eventually, we must approximate this infinite delay line with one of finite extent. This truncation should throw away only taps of small magnitude, and thus the tap weights should decrease in magnitude away from the center tap. Consequently, we demand that $\alpha_j \to 0$ as $|j| \to \infty$. Thus,

$$\begin{aligned} C_1 &= C_2 = \cdots = C_I = 0 \quad \text{for} \quad j \text{ negative.} \\ D_1 &= D_2 = \cdots = D_\Omega = 0 \quad \text{for} \quad j \text{ positive.} \end{aligned} \tag{18}$$

In effect, we have two solutions; one for taps with negative subscripts and one for taps with positive subscripts.

$$\begin{aligned} \alpha_j &= C_1\theta_1^j + C_2\theta_2^j + \cdots + C_I\theta_I^j \quad \text{for} \quad j > 0 \\ \alpha_j &= D_1\phi_1^j + D_2\phi_2^j + \cdots + D_\Omega\phi_\Omega^j \quad \text{for} \quad j < 0. \end{aligned} \tag{19}$$

The region of overlap in equation (13), that is, the region where the equations involve tap weights with both positive and negative subscripts, will determine the arbitrary constants. To illustrate this, let us consider the impulse response of Fig. 2. For one set (No. 8) of samples, the samples are -2, 3, 11, -6 and the roots are 0.5, 3, -2. Thus

$$\begin{aligned} \alpha_j &= C_1(0.5)^j \qquad\qquad\quad \text{for} \quad j > 0 \\ \alpha_j &= D_1(3)^j + D_2(-2)^j \quad \text{for} \quad j < 0. \end{aligned} \tag{20}$$

Now, C_1, D_1, and D_2 may be found by using equation (20) in equation (13). However, first we must assign an origin in our pulse. Tbat is, we must decide whether -2, or 3, or 11, or -6 is to be called β_0. Solutions exist for each of these possibilities, but only one of these has certain desirable properties which we will discuss later. In this case, β_0 should be the sample of magnitude 11. With this assignment, equation (13) may be written

$$\vdots$$

$$-2[C_1(.5)^4] + 3[C_1(.5)^3] + 11[C_1(.5)^2] - 6[C_1(.5)] = 0 \qquad (21a)$$

$$-2[C_1(.5)^3] + 3[C_1(.5)^2] + 11[C_1(.5)] - 6\alpha_0 = 0 \qquad (21b)$$

$$-2[C_1(.5)^2] + 3[C_1(.5)] + 11\alpha_0 - 6[D_1(3)^{-1} + D_2(-2)^{-1}] = 1 \qquad (21c)$$

$$\begin{aligned} -2[C_1(.5)] + 3\alpha_0 &+ 11[D_1(3)^{-1} + D_2(-2)^{-1}] \\ &- 6[D_1(3)^{-2} + D_2(-2)^{-2}] = 0 \end{aligned} \qquad (21d)$$

$$\begin{aligned} -2\alpha_0 + 3[D_1(3)^{-1} + D_2(-2)^{-1}] &+ 11[D_1(3)^{-2} + D_2(-2)^{-2}] \\ &- 6[D_1(3)^{-3} + D_2(-2)^{-3}] = 0 \end{aligned} \qquad (21e)$$

$$\begin{aligned} &-2[D_1(3)^{-1} + D_2(-2)^{-1}] + 3[D_1(3)^{-2} + D_2(-2)^{-2}] \\ &\quad + 11[D_1(3)^{-3} + D_2(-2)^{-3}] - 6[D_1(3)^{-4} + D_2(-2)^{-4}] = 0 \end{aligned} \qquad (21f)$$

$$\vdots$$

Equations (21a) and (21f) and all others above and below these two are automatically satisfied for any choice of C_1, D_1, D_2. Consider equation (21b). In order for it to be satisfied, α_0 must be equal to C_1 $(0.5)^0$. Similarly, for equation (21e) to be satisfied, α_0 must equal $D_1(3)^0 + D_2(-2)^0$,

$$C_1 = D_1 + D_2 = \alpha_0 \qquad (22)$$

and equation (21) yields the following values for the C and D constants.

$$\left.\begin{aligned} D_1 &= 2/95 \\ D_2 &= 5/95 \\ C_1 &= 7/95 \end{aligned}\right\} \beta_0 = 11. \qquad (23)$$

If we had chosen to let β_0 be the sample of magnitude 3, we would obtain

$$\left.\begin{aligned} D_1 &= -10/95 \\ D_2 &= 15/95 \\ C_1 &= 5/95 \end{aligned}\right\} \beta_0 = 3. \tag{24}$$

If the sample of magnitude 2 is made β_0, then

$$\left.\begin{aligned} D_1 &= 9/25 \\ D_2 &= 4/25 \\ C_1 &= 1/50 \end{aligned}\right\} \beta_0 = -2. \tag{25}$$

In these last two examples, the solutions may be considered inferior because the tap weights away from the center tap will be larger than in the first example. However, stronger objections to the last two choices for β_0 will be raised shortly.

A point of considerable interest is apparent in the development above. If any of the roots fall on the unit circle, then no solutions exist in which the tap weights decay in magnitude in both directions away from the center tap.

IV. TRUNCATION EFFECTS

In any practical equalizer, the number of taps available is not infinite. Thus, we must investigate the effects of limiting tbe number of taps to some reasonable finite value. For example, let us suppose that we have $N + n + 1$ taps.

$$\alpha_j = 0 \quad \text{for} \quad \begin{cases} j < -n \\ j > N \end{cases}. \tag{26}$$

Let us consider two schemes for setting the truncted tap weights. A more or less obvious way is simply to take the infinite solution [for example, equation (23)] for all available taps. This can be represented in matrix form:

$$[B_\infty][\alpha_T] = [\gamma]. \tag{27}$$

The matrix $[B_\infty]$ is the infinite matrix of β values shown in equation (13). The truncated tap set is $[\alpha_T]$,

$$[\alpha_T] = \begin{bmatrix} \vdots \\ 0 \\ 0 \\ \alpha_N \\ \alpha_{N-1} \\ \alpha_1 \\ \alpha_0 \\ \alpha_{-1} \\ \vdots \\ \alpha_{-n} \\ 0 \\ 0 \\ \vdots \end{bmatrix} \tag{28}$$

and $[\gamma]$ is the matrix of output sample values. Notice that $[\alpha_T]$ may be written as

$$[\alpha_T] = [\alpha_\infty] - [\alpha_t] \tag{29}$$

where $[\alpha_\infty]$ is the infinite set of tap weights which give us the desired output and

$$[\alpha_t] = \begin{bmatrix} \vdots \\ \alpha_{N+2} \\ \alpha_{N+1} \\ 0 \\ 0 \\ 0 \\ \vdots \\ 0 \\ \alpha_{-n-1} \\ \alpha_{-n-2} \\ \vdots \end{bmatrix}. \tag{30}$$

Thus, equation (27) becomes

$$[B_\infty][\alpha_T] = [B_\infty][\alpha_\infty] - [B_\infty][\alpha_t] = [\mu] - [\epsilon]. \tag{31}$$

The desired output $[\mu]$, all zero samples except the central one with unit magnitude, is given by $[B_\infty][\alpha_\infty]$ while the error set $[\epsilon]$ is given by $[B_\infty][\alpha_t]$.

It is possible to calculate this error set exactly, but an approximate bound should suffice. Assume that N and n are large so that the α values in $[\alpha_t]$ depend only on the roots which are the nearest to the unit circle. Let the magnitude of nearest inside and outside roots be θ_u and ϕ_u. Then approximately

$$\alpha_j \cong \begin{matrix} \theta_u^j & \text{for} & j > N \\ \phi_u^j & \text{for} & j < -n \end{matrix} . \tag{32}$$

Consequently the largest component of any error term in $[\epsilon]$ will be about

$$\text{Max comp} \cong \beta_{\max} \cdot C_u \theta_u^{N+1} \cong \beta_{\max}\ D_u \phi_u^{-n-1} \tag{33}$$

and each error term is the sum of less than $N + n + 1$ components. Thus, an upper bound on the error terms is

$$|\epsilon_{\max}| < |(N + n) C_u \theta_u^{N+1}| \cong |(N + n)\ D_u \phi_u^{-n-1}| \tag{34}$$

which vanishes as N and n are made very large.

V. A SECOND METHOD OF TRUNCATION

This second method offers no improvement in ultimate equalization over the method just discussed. However, it does lend itself to iterative adjustment techniques whereas the first method tacitly assumes a computation which provides the proper infinite solution to begin with. In this second method we require that all the output samples (excluding γ_0) corresponding to the $N + n + 1$ taps, that is, $\gamma_N, \gamma_{N-1}, \cdots, \gamma_2, \gamma_1, \gamma_{-1}, \gamma_{-2}, \cdots, \gamma_{-n+1}, \gamma_{-n}$, be zero. This criterion may be called the Lucky criterion since it is the one R. W. Lucky has used in his work.[5]

What does this criterion mean in terms of the solutions (powers of impulse response roots) we discussed for the infinite tap case? We are essentially constraining our system further by another set of boundary conditions. We will call these boundaries the positive boundary at α_N and the negative boundary at α_{-n}, in addition to the central boundary around α_0 where we have already discussed satisfying boundary con-

ditions (those specifying $C_1, C_2, \ldots, D_1, D_2, \ldots$). As an example, let us consider the equations at the positive boundary.

$$\beta_0\alpha_N + \beta_{-1}\alpha_{N-1} + \cdots \cdots + \beta_{-m}\alpha_{N-m} = \gamma_N = 0 \quad (35a)$$

$$\beta_1\alpha_N + \beta_0\alpha_{N-1} + \cdots + \beta_{-m}\alpha_{N-(m+1)} = \gamma_{N-1} = 0 \quad (35b)$$

$$\vdots \qquad\qquad \vdots$$

$$\beta_{M-1}\alpha_N + \beta_{M-2}\alpha_{N-1} + \cdots + \beta_{-m}\alpha_{N-(m+M-1)} = \gamma_{N-(M-1)} = 0 \quad (35c)$$

$$\beta_M\alpha_N + \beta_{M-1}\alpha_{N-1} + \cdots \cdots + \beta_{-m}\alpha_{N-(m+M)} = \gamma_{N-M} = 0 \quad (35d)$$

Equation (35d) can be satisfied with the solutions determined by the central boundary, that is

$$\alpha_i = C_1\theta_1^i + C_2\theta_2^i + \cdots + C_I\theta_I^i . \quad (36)$$

However, this solution will not satisfy the M equations above equation (35d) since the complete set of β values is missing in these equations. The exponentially growing solutions which were discarded earlier must be used now. Thus,

$$\alpha_j = C_1\theta_1^j + \cdots + C_I\theta_I^j + c_1\theta_1^j + c_2\theta_2^j + \cdots + c_\Omega\phi_\Omega^j \quad \text{for} \quad j > 0 \quad (37)$$

$$\alpha_j = D_1\phi_1^j + \cdots + D_\Omega\phi_\Omega^j + d_1\theta_1^j + d_2\theta_2^j + \cdots + d_I\theta_I^j \quad \text{for} \quad j < 0 \quad (38)$$

must be used in order to satisfy all the boundary equations. If $\Omega = M$, then the lower case c's in equation (37) provide just enough constants to satisfy the M equations of equations (35a through c). Furthermore, I will equal m and the I lower case d's will provide just enough constants to satisfy the m boundary equations at the negative boundary.

To illustrate the preceding discussion, let us return to the specific example discussed previously. We use the results of equation (23) with

$$\beta_2 = -2, \quad \beta_1 = 3, \quad \beta_0 = 11, \quad \beta_{-1} = -6. \quad (39)$$

The positive boundary equations are

$$\begin{aligned} &11[C_1(.5)^N + c_1(3)^N + c_2(-2)^N] \\ &\qquad - 6[C_1(.5)^{N-1} + c_1(3)^{N-1} + c_2(-2)^{N-1}] = 0 \\ &3[C_1(.5)^N + c_1(3)^N + c_2(-2)^N] + 11[C_1(.5)^{N-1} + c_2(-2)^{N-1}] \\ &\qquad - 6[C_1(.5)^{N-2} + c_1(3)^{N-2} + c_2(-2)^{N-2}] = 0 \end{aligned} \quad (40)$$

$$\begin{bmatrix} 3^{N-1}(33 - 6) & -2^{N-1}(-22 - 6) \\ 3^{N-2}(27 + 33 - 6) & -2^{N-2}(12 - 22 - 6) \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \end{bmatrix}$$

$$= \begin{bmatrix} .5^N(11 - 12)C_1 \\ .5^N(3 + 22 - 24)C_1 \end{bmatrix}. \tag{41}$$

If we assume that N is large enough so that the value of the C and D variables are unaffected by inclusion of the other roots, then these C and D values may be used to determine the c and d values. With $C_1 = 7/95$, equation (41) becomes

$$\begin{bmatrix} 9 & 14 \\ 6 & -4 \end{bmatrix} \begin{bmatrix} c_1' \\ c_2' \end{bmatrix} = \begin{bmatrix} -\mathcal{C} \\ \mathcal{C} \end{bmatrix} \tag{42}$$

where

$$c_1' = 3^N c_1 \qquad c_2' = (-2)^N c_2 \qquad \mathcal{C} = \left(\frac{7}{95}\right)(.5)^N. \tag{43}$$

Thus,

$$\begin{aligned} c_1' &= \mathcal{C}/12 & c_2' &= -\mathcal{C}/8 \\ c_1 &= \frac{\left(\frac{7}{95}\right)(.5)^N}{(3)^N} & c_2 &= \frac{\left(\frac{7}{95}\right)(.5)^N}{(-2)^N}. \end{aligned} \tag{44}$$

Similarly, d_1 can be found by the single boundary equation at the negative boundary.

$$d_1 = -\left[\frac{2\left(\frac{2}{95}\right)(3)^{-n} + 9\left(\frac{5}{95}\right)(-2)^{-n}}{12}\right](.5)^n. \tag{45}$$

The results worked out above can be roughly represented graphically as in Fig. 5 where the magnitudes of the roots to the tap-number power are illustrated. This figure shows what will be called a "good" solution. That is, the decaying solutions predominate with the growing solutions contributing only a small amount at the positive and negative boundaries. The residual errors, that is, the γ_k values for $k > N$ and $k < -n$, will be of the same order of magnitude as those of the first truncation method.

In order to have a good solution as demonstrated above, the proper

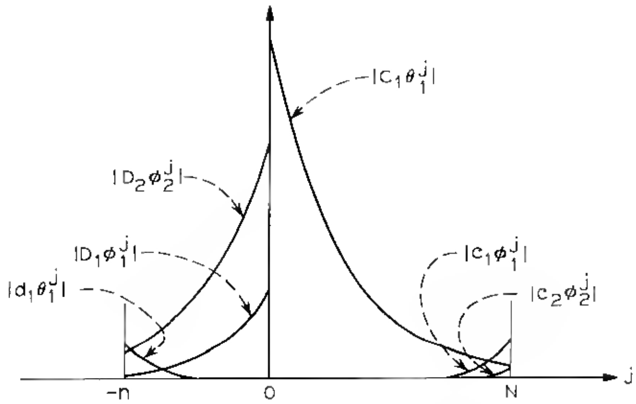

Fig. 5 — Solution behavior for truncated equalizer.

number of arbitrary constants must be available to satisfy the positive and negative boundary equations. The number of equations at the positive boundary is M, and the number of available constants (outside roots) is Ω; similarly, there are I constants (inside roots) available for the m equations at the negative boundary. Thus, the necessary conditions for a good solution are

$$\left.\begin{array}{l}\text{no. outside roots} = \Omega = M = \text{no. of samples following } \beta_0 \\ \text{no. inside roots} = I = m = \text{no. of samples preceding } \beta_0\end{array}\right\}. \qquad (46)$$

We may use the same simple pulse to demonstrate a "bad" solution. If we go back to equation (24), we see that an infinite (nontruncated) solution exists for the situation in which we decided to let $\beta_0 = 3$. The number of inside and outside roots remain the same in this case, but M and m are both changed. Now $M = 1$ and $m = 2$, and equations (46) are no longer satisfied. All the boundary conditions can still be satisfied, but not in such a simple manner. That is to say, in the preceding example, satisfying equations (46), a separation of solutions is possible. The central boundary specifies the values of the upper-case constants, then the lower-case constants are set to compensate for the truncation effect at the positive and negative boundaries. Since the necessary compensation is small, and since the effect of the lower-case constants dies out towards the central boundary, only minor or negligible corrections to the upper-case constants are necessary to keep the central boundary conditions satisfied.

Now let us consider what takes place when this step-by-step solu-

tion is attempted when equations (46) are not satisfied. Suppose C_1, D_1, and D_2 are set to satisfy the central boundary according to equation (24). At the positive boundary, only one of the lower-case c's is needed so no difficulty arises and the positive tap solutions will, to this point, be little different from that illustrated in Fig. 5. However, at the negative boundary, the single available arbitrary constant d_1 is not enough to satisfy the two boundary equations which occur in this case. All the upper and lower case d-constants can be adjusted to satisfy the negative boundary conditions, but this will destroy the equilibrium of the central boundary solution since any change in the upper-case D's does effect the central equations. If one of the upper-case D's is constrained by the negative boundary, then the remaining unused lower-case c can be brought in to provide enough arbitrary constants to satisfy the central boundary conditions. The net effect of all this will generally be that a growing solution must be made to have a nonnegligible contribution at the central boundary. Consequently, it will be large at the positive boundary. This is illustrated in Fig. 6.

When the situation discussed above and represented in Fig. 6 occurs, the residual values of γ_k outside the equalization region will be large and will generally grow larger as the number of taps is increased.

As an actual example of a bad solution of the type discussed above, consider the pulse illustrated on page 563 of Ref. 5. The polynomial

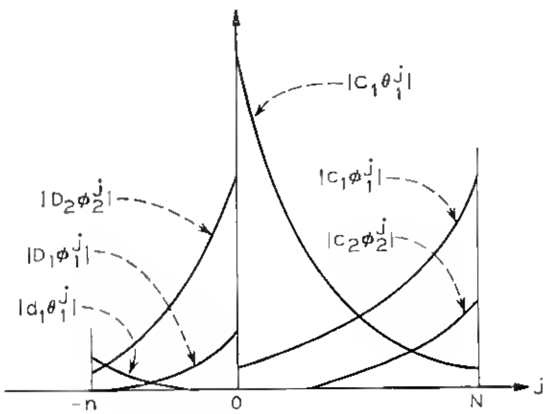

Fig. 6 — Typical solution behavior when the conditions of equations (46) are not satisfied.

which represents this pulse with β_o set as the peak value is

$$B(z) = 15z^8 - 20z^7 + 40z^6 - 60z^5 + 0z^4 + 115z^3 + 60z^2 + 20z + 10. \tag{47}$$

Notice that with this choice for β_o,

$$M = 5 \quad m = 3. \tag{48}$$

The root locations for this polynomial are illustrated in Fig. 7. There are four inside roots ($I = 4$) and four outside roots ($\Omega = 4$). Consequently, equations (46) are not satisfied and no good truncated solution exists. This is verified by the fact that when Lucky attempted to equalize this particular pulse, the equalizer gave an output pulse with large γ_k, $k > N$. This indicates that a solution such as that sbown in Fig. 6 has been approached.

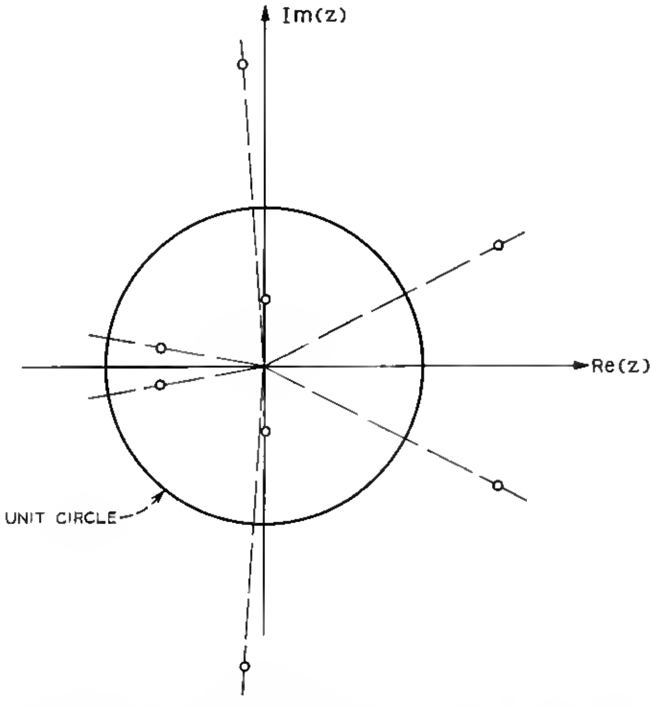

Fig. 7 — Root locations of pulse page 563 of Ref. 5. $B(z) = 15z^8 - 20z^7 + 40z^6 - 60z^5 + 0z^4 + 115z^3 + 60z^2 + 20z + 10$ $\theta_1, \theta_1^* = 0.404 \angle \pm 89.8°$ $\theta_2, \theta_2^* = 0.652 \angle \pm 169.9°$ $\phi_1, \phi_1^* = 1.642 \angle \pm 27.7°$ $\phi_2, \phi_2^* = 1.894 \angle \pm 94.4°$.

If β_o were shifted (delayed) one sample so that it would be the coefficient of the fourth-degree term of $B(z)$, then equations (46) would be satisfied and a good truncated solution would exist. Notice however, that in this case β_o would be zero which presents a severe problem in iteratively searching for this good solution.

VI. ITERATIVE SEARCH FOR EQUALIZATION

The discussion has indicated that, except for singular cases where one or more roots of $B(z)$ are on the unit circle, equalization solutions exist (for the infinite tap case) in which tap weights decrease exponentially in both directions from the center tap. Furthermore, good truncated solutions exist which force all γ_k $(-n \leqq k \leqq N,\ k \neq 0)$ to zero if β_0 is selected to satisfy equations (46), which can always be done. The question which now arises is whether a simple iterative search routine will lead to a desirable equalization.

When presented with an impulse response to be equalized using a truncated equalizer and the Lucky criterion of forcing output samples to zero, a hierarchy of questions must be considered:

(*i*) Is the pulse equalizable? That is, are all the roots off the unit circle? (If there are roots on or very close to the unit circle, a change of timing, that is, varying τ in equation (3), will usually move the roots off the unit circle.)

(*ii*) If the pulse is equalizable, does a good truncated solution exist? A shifting of β subscripts can always guarantee the existence of a good solution by satisfying equations (46), but will sometimes create convergence problems.

(*iii*) If the pulse is equalizable and a good truncated solution exists, will a simple iterative search find this solution?

The iterative method of searching for a solution which we consider first consists of measuring the value (γ_k) of the kth output sample, then subtracting some part of this from the kth tap weight.* That is

$$\alpha_k^{(r+1)} = \alpha_k^{(r)} - \Delta\gamma_k^{(r)} \tag{49}$$

where the superscripts indicate the iteration number and Δ is a positive number less than one. This iterative process is not identical to the method presented in Ref. 5 which increments the tap according

* This is not the only possible iterative search method, but it is one of the simplest.

to only the sign of γ_k.

$$\alpha_k^{(r+1)} = \alpha_k^{(r)} - \Delta \operatorname{sgn} [\gamma_k^{(r)}]. \tag{50}$$

However, the two methods are very similar, and convergence of one will almost always guarantee convergence of the other.

In matrix notation, the linear iterative search of equation (49) can be written as

$$\alpha^{(r+1)} = \alpha^{(r)} - \Delta\epsilon^{(r)} \tag{51}$$

$$\epsilon^{(r)} = B\alpha^{(r)} - \mu \tag{52}$$

where

$$\alpha = \begin{bmatrix} \alpha_N \\ \alpha_{N-1} \\ \vdots \\ \alpha_1 \\ \alpha_0 \\ \alpha_{-1} \\ \vdots \\ \alpha_{-n} \end{bmatrix} \qquad \mu = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}.$$

$$B = \begin{bmatrix} \beta_0 & \beta_{-1} & \beta_{-2} & \cdots\cdots & \beta_{-m} & 0 & 0 & 0 & \cdots & 0 \\ \beta_1 & \beta_0 & \beta_{-1} & \beta_{-2} & \cdots\cdots & \beta_{-m} & & & & \\ \beta_2 & \beta_1 & \beta_0 & \beta_{-1} & \beta_{-2} & \cdots\cdots & \beta_{-m} & & & \\ \cdots & \beta_2 & \beta_1 & \beta_0 & \beta_{-1} & \beta_{-2} & \cdots\cdots & \beta_{-m} & & \\ & & & & \beta_0 & & & & & \\ & & & & & \beta_0 & & & & \\ & & & & & & \beta_0 & & & \\ & & & & & & & \cdot & & \\ & & & & & & & & \cdot & \\ & & & & & & & & & \beta_0 \end{bmatrix} \tag{53}$$

Equation (52) can be modified by using equation (51),

$$\begin{aligned}\epsilon^{(r)} &= B[\alpha^{(r-1)} - \Delta\epsilon^{(r-1)}] - \mu \\ &= (I - \Delta B)\epsilon^{(r-1)} \\ &= (I - \Delta B)^r \epsilon^{(0)} \end{aligned} \tag{54}$$

where I is the identity matrix.

Clearly, the iterative routine will converge if

$$\lim_{r\to\infty} \epsilon^{(r)} = 0. \tag{55}$$

However, it does not quarantee convergence to a good solution as illustrated by the pulse of equation (47) and Fig. 7. There, convergence does occur as was illustrated in Ref. 5 and hence, equation (55) is satisfied but convergence is to a bad solution. In that case no good solution exists to which the routine can converge.

The necessary and sufficient condition for convergence is that

$$\lim_{r\to\infty} (I - \Delta B)^r = 0. \tag{56}$$

The matrix of equation (56), that is, the matrix $I - \Delta B$ will converge to zero if and only if all its eigenvalues are less than one in magnitude.[6] This is equivalent to the condition, illustrated in Fig. 8, that the eigenvalues of B lie within a circle of radius $1/\Delta$ centered at $1/\Delta$ on the real axis. Assuming that Δ can be made as small as necessary, *a necessary and sufficient condition for iterative convergence is that all*

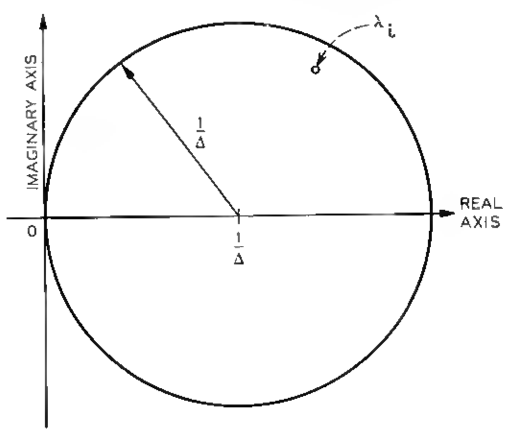

Fig. 8 — Necessary relation of Δ to location of eigenvalue in complex plane.

eigenvalues of B should have a positive real part. If this condition is not satisfied, $(I-\Delta B)^r$ will diverge, and even if the initial error vector $\epsilon^{(0)}$ is very small, the final error vector will grow without bound.

The italicized condition above, coupled with the root location condition of equation (46), guarantees convergence to a good solution. Certain other more restrictive sufficient (but not necessary) conditions can be derived. These conditions may be easier to check for a channel impulse response.

6.1 *Necessary and Sufficient Condition for Monotonic Convergence.*

The convergence of the iterative process will be called monotonic if for any starting error vector $\epsilon^{(0)}$, the following inequalities hold (a superscript t indicates the transpose of a matrix).

$$\| \epsilon^{(0)} \| > \| \epsilon^{(1)} \| > \| \epsilon^{(2)} \| > \cdots > \| \epsilon^{(r)} \| > \| \epsilon^{(r+1)} \| > \cdots \tag{57}$$

$$\| \epsilon^{(r)} \| = \epsilon^{(r)t} \epsilon^{(r)} = \sum_{i=-n}^{N} [\epsilon_i^{(r)}]^2. \tag{58}$$

Since

$$\epsilon^{(r+1)} = (I - \Delta B)\epsilon^{(r)}, \tag{59}$$

it follows that

$$\begin{aligned} \| \epsilon^{(r+1)} \| &= \epsilon^{(r+1)t}\epsilon^{(r+1)} = \epsilon^{(r)t}(I - \Delta B)^t (I - \Delta B)\epsilon^{(r)} \\ &= \epsilon^{(r)t}\epsilon^{(r)} + \Delta^2 [B\epsilon^{(r)}]^t [B\epsilon^{(r)}] - \Delta\epsilon^{(r)t}[B^t + B]\epsilon^{(r)} \\ &= \| \epsilon^{(r)} \| + \Delta^2 \| B\epsilon^{(r)} \| - \Delta\epsilon^{(r)t}[B^t + B]\epsilon^{(r)}. \end{aligned} \tag{60}$$

Thus, equation (57) will be satisfied for all possible initial error vectors if an only if

$$\epsilon^t [B^t + B]\epsilon > \Delta \| B\epsilon \| = \Delta (B\epsilon)^t (B\epsilon). \tag{61}$$

Since the right side of the inequality is always positive, the inequality can be satisfied for all possible nonzero ϵ only if $[B^t + B]$ is positive-definite. If this is true and

$$\min_{\epsilon} \frac{\epsilon^t [B^t + B]\epsilon}{\| B\epsilon \|} > \Delta \tag{62}$$

tben equation (61) will be satisfied for all ϵ and the iterative process will be monotonically convergent.

If $[B^t + B]$ is not positive-definite, convergence can still occur. However, stating the conditions for convergence becomes more difficult. For example, if $[B^t+B]$ is not positive-definite then for some vector, say $\epsilon^{(r-1)}$,

$$\epsilon^{(r-1)t}[B^t + B]\epsilon^{(r-1)} < 0 \tag{63}$$

and

$$\| \epsilon^{(r-1)} \| < \| \epsilon^{(r)} \| . \tag{64}$$

But if for $\epsilon = \epsilon^{(r-1)}$

$$\begin{aligned}(B\epsilon)^t[B^t + B](B\epsilon) > \frac{1}{2^3} \{ & \Delta^4 \| B^2\epsilon \| + 2\Delta^2 \| B\epsilon \| \\ & + \Delta^2 \| [B^t + B]\epsilon \| - 2\Delta\epsilon^t[B^t + B]\epsilon\}\end{aligned} \tag{65}$$

then

$$\| \epsilon^{(r-1)} \| > \| \epsilon^{(r+1)} \| . \tag{66}$$

Very roughly speaking, equations such as (65) which can be developed indicate that $[B^t+B]$ should yield a predominantly positive quadratic form in order to have convergence. What we mean by this can best be illustrated graphically, as in Fig. 9.

6.2 *Further Sufficient Conditions for Monotonic Convergence*

The positive-definiteness of $[B^t+B]$ is a necessary and sufficient condition for monotonic convergence. Being somewhat more restrictive will yield other sufficient (but not necessary) conditions. We notice that $[B^t + B]$ is a Toeplitz matrix, that is,

$$[B^t + B] = [b_{ij}], \quad b_{ij} = b_{|i-j|} = \beta_{(i-j)} + \beta_{(j-i)} . \tag{67}$$

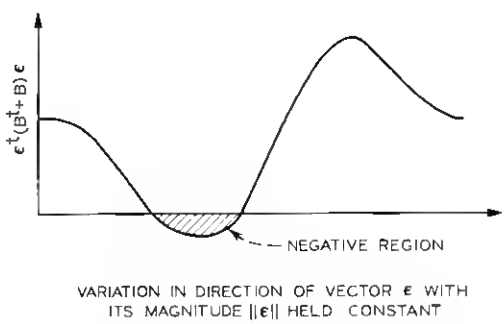

Fig. 9 — Illustrating a predominately positive quadratic form.

Thus, it has some of the necessary attributes of a correlation matrix. A correlation matrix is positive-definite, so if we can show that $[B^t+B]$ is a correlation matrix, then this is sufficient for positive-definiteness. If $[B^t + B]$ is a correlation matrix then the polynomial $B(z) + B^t(z) = P(z)$, where

$$B(z) = \beta_M z^{M+m} + \cdots + \beta_{-m}$$
$$B^t(z) = \beta_{-m} z^{2m} + \cdots + \beta_M z^{m-M} \tag{68}$$

must have no odd order roots on the unit circle. Root locus methods may be applied here, that is, the loci of

$$\frac{B(z)}{B^t(z)} = -k \tag{69}$$

must not cross the unit circle. Notice that this is a necessary condition that $[B^t + B]$ be a correlation matrix and it is sufficient but *not* necessary for $[B^t + B]$ being positive-definite.

A more restrictive condition leads to another sufficiency condition. If (Lucky condition, Ref. 5)

$$\beta_0 > \sum_{k \neq 0} |\beta_k|, \quad \text{that is,} \quad D_0 < 1 \tag{70}$$

then it is easy to see that $B^t(z) + B(z)$ has no roots on the unit circle and thus $[B^t + B]$ is a correlation matrix and positive-definite. Notice that equation (70) is not a necessary condition for $[B^t + B]$ being a correlation matrix.

The various conditions discussed above are summarized in Fig. 10.

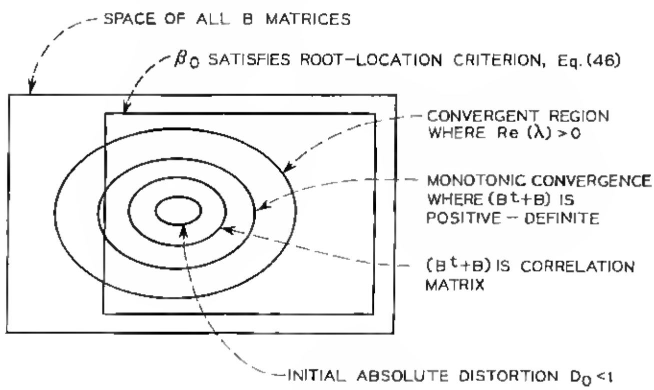

Fig. 10 — Illustrating the different convergence regions.

VII. MODIFIED ITERATION PROCESSES

In the sample iteration process which has been discussed to this point, the change in tap weight α_j depends only on the output sample value γ_j. Let us generalize so that the change depends on a linear combination of output samples.

$$\left.\begin{aligned} B\alpha^{(r)} &= \gamma^{(r)} \\ \epsilon^{(r)} &= \gamma^{(r)} - \mu = B\alpha^{(r)} - \mu \\ \alpha^{(r+1)} &= \alpha^{(r)} - \Delta V\epsilon^{(r)} \end{aligned}\right\}. \tag{71}$$

That is to say, the $(r + 1)$st value of tap j is

$$\alpha_j^{(r+1)} = \alpha_j^{(r)} - \Delta \sum_k v_{jk}\epsilon_k^{(r)}. \tag{72}$$

Manipulation of equations (71) leads to a result similar to equation (54)

$$\epsilon^{(r)} = [I - \Delta BV]^r \epsilon^{(0)}. \tag{73}$$

Consequently, all the sufficient and necessary conditions which have been developed on the previous pages can now be applied to the matrix BV. Now, however, we have considerably more latitude since we are free to specify V.

As an example, let us suppose that V is chosen to equal B^t. Then

$$BV = BB^t = \text{positive-definite} \tag{74}$$

and monotonic convergence is guaranteed. This is a particularly appealing way of selecting V since the sample weighting can be determined directly from the initial channel impulse response.

$$v_{jk} = v_{j-k} = \beta_{k-j} . \tag{75}$$

It is very interesting that the weighting suggested above is very nearly equivalent to inserting a tapped-delay-line matched filter ahead of the equalizer. A matched filter, whose tap weights are equal to the β values in reverse order, will yield an output whose samples will form a B matrix which is a correlation matrix. Thus, the iterative search will be monotonically convergent in this case also. The weighting suggested above yields the same matrix except for "edge effects."

This can be illustrated by the following example. Suppose there are just three β values: $\beta_{-1} = 1$, $\beta_0 = 2$, $\beta_1 = -2$. Then, with V set

equal to B^t, and assuming a six-tap equalizer so that B is a 6×6 matrix, we obtain

$$BV = BB^t = \begin{bmatrix} 5 & -2 & -2 & 0 & 0 & 0 \\ -2 & 9 & -2 & -2 & 0 & 0 \\ -2 & -2 & 9 & -2 & -2 & 0 \\ 0 & -2 & -2 & 9 & -2 & -2 \\ 0 & 0 & -2 & -2 & 9 & -2 \\ 0 & 0 & 0 & -2 & -2 & 8 \end{bmatrix}. \tag{76}$$

On the other hand, the modified B matrix following a matched filter would be

$$B_m = \begin{bmatrix} 9 & -2 & -2 & 0 & 0 & 0 \\ -2 & 9 & -2 & -2 & 0 & 0 \\ -2 & -2 & 9 & -2 & -2 & 0 \\ 0 & -2 & -2 & 9 & -2 & -2 \\ 0 & 0 & -2 & -2 & 9 & -2 \\ 0 & 0 & 0 & -2 & -2 & 9 \end{bmatrix} \tag{77}$$

Setting $V = B^t$ will guarantee monotonic convergence. However, since in most cases B is such that $[B^t + B]$ is close to being positive-definite, it is probable that a less extensive V would be sufficient to guarantee monotonic convergence. As an example, suppose that V is chosen to be a small deviation on the standard iteration of equation (58)

$$V = I + \delta B^t, \qquad \delta > 0. \tag{78}$$

Now,

$$BV = B + \delta BB^t \tag{79}$$

will have a quadratic form which is greater for every vector ϵ than the quadratic form for the matrix B alone. Thus, if the negative region such as is illustrated in Fig. 8 is small, then BV can become positive-definite for relatively small δ.

Perhaps a more reasonable way of selecting a V which approximates B^t is to modify equation (75) in the following manner.

$$v_{ik} = v_{i-k} = \begin{cases} \beta_{i-k} & \text{if} \quad | \beta_{k-i} | \geqq L \\ 0 & \text{if} \quad | \beta_{k-i} | < L \end{cases}. \tag{80}$$

Thus, only the more significant values of β are used in weighting the errors. In general such a V will force B towards positive-definiteness. However, there seems to be no general rule for selecting the critical value of L, that is the maximum L which will just permit positive-definiteness.

VIII. CONCLUDING REMARKS

Virtually all input pulses are equalizable in the sense that there exist tap-weight adjustments which will force the output samples in the adjustment interval to zero while the output samples out of the adjustment interval remain small. Furthermore, the residual samples outside of the adjustment interval will become smaller as the number of taps (length of transversal filter) is increased.

There are just two necessary conditions in order for the preceding statement to hold. The first is that the polynomial representing the input pulse have no roots on the unit circle. Although the singular case where roots are exactly on the unit circle is highly improbable, roots very near the unit circle lead to relatively larger residual errors and greater potential for instability.

The second necessary condition is that the selection of the central sample value must be such that equation (46) is satisfied.

Although a pulse may be equalizable, the simple first-order iterative search for the proper tap weights given by equations (51) and (52) may not be convergent. If it is convergent, and assuming the two conditions above are satisfied, it will converge to the proper tap-weight settings. If it is not convergent, it will be divergent with increasing errors in the adjustment interval. The convergence or divergence is independent of the initial tap settings. Thus, even though the tap weights might be set to optimum initially, if the system is in the iterative search mode and is divergent, it will eventually diverge.

The necessary and sufficient condition for convergence is given in Section VI along with a hierarchy of more stringent sufficient conditions. In general, convergence will be dependent upon the absolute timing of the sampling. Consequently, a particular pulse which is equalizable for two different timings might be convergent for one timing and divergent for the other.

If the first-order iterative procedure is divergent, a more complex

weighting of the output errors in adjusting the tap weights can improve the situation. At least one weighting given by equation (75) will guarantee convergence.

REFERENCES

1. Nyquist, H., "Certain Topics in Telegraph Transmission Theory," AIEE Trans. Commun. and Elec., *47*, No. 2 (April 1928), pp. 617–644.
2. Bennett, W. R. and Davey, J. R., *Data Transmission*, New York: McGraw-Hill, 1965, p. 61.
3. Gibby, R. A. and Smith, J. W., "Some Extensions of Nyquist's Telegraph Transmission Theory," B.S.T.J., *46*, No. 7 (September 1965), pp. 1487–1510.
4. Lucky, R. W., "Techniques for Adaptive Equalization of Digital Communication Systems," B.S.T.J., *45*, No. 2 (February 1966), pp. 255–286.
5. Lucky, R. W., "Automatic Equalization for Digital Communication," B.S.T.J., *44*, No. 4 (April 1965), pp. 547–588.
6. Householder, A. S., *The Theory of Matrices in Numerical Analysis*, New York: Blaisdell Publishing Co., 1964, pp. 53–55.